

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-089786

(43)Date of publication of application : 31.03.2000

(51)Int.Cl.

G10L 15/22

(21)Application number : 10-254377

(71)Applicant : NIPPON HOSO KYOKAI <NHK>

(22)Date of filing : 08.09.1998

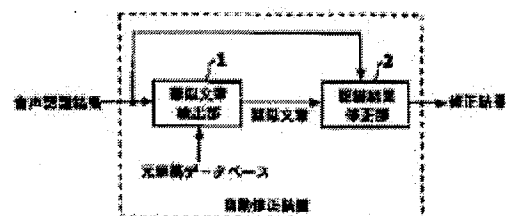
(72)Inventor : ONOE KAZUHO
IMAI TORU
ANDO AKIO

(54) METHOD FOR CORRECTING SPEECH RECOGNITION RESULT AND APPARATUS THEREFOR

(57)Abstract:

PROBLEM TO BE SOLVED: To lessen the manipulation labor of a check person by comparing a second manuscript before final determination and a speech recognition result by an information processor, detecting the mismatch portion of the second manuscript and the speech recognition result and correcting the speech recognition result by substituting the detected mismatch portion of the second manuscript side with the mismatch portion of the speech recognition result.

SOLUTION: A similar sentence detecting section 1 collates the speech recognition result and the respective sentences of the original manuscript in word units and determines a collation score in a collation score calculating section. Next, if the collation score is above the threshold, this original manuscript is regarded to be the similar sentence of the speech recognition result in a similar sentence decision section. In a speech recognition correction section 2, the section, where the speech recognition result and the word of the similar sentence do not coincide, is detected when the presence of the similar sentence in the similar sentence detecting section 1 is decided. Next, if the differences between the speech recognition result of the mismatch section and the respective word numbers of the similar sentences are below a certain specified number the word of the speech recognition result is substituted with the word of the similar sentence.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-89786

(P2000-89786A)

(43) 公開日 平成12年3月31日 (2000.3.31)

(51) Int.Cl.⁷

識別記号

F I

メモ* (参考)

G 1 0 L 15/22

G 1 0 L 3/00

5 6 1 E

審査請求 未請求 請求項の数14 O L (全 9 頁)

(21) 出願番号 特願平10-254377

(22) 出願日 平成10年9月8日 (1998.9.8)

特許法第30条第1項適用申請有り 平成10年3月17日
社団法人日本音響学会発行の「日本音響学会平成10年度
春季研究発表会講演論文集 ▲I▼」に発表

(71) 出願人 000004352

日本放送協会

東京都渋谷区神南2丁目2番1号

(72) 発明者 尾上 和穂

東京都世田谷区砦一丁目10番11号 日本放
送協会 放送技術研究所内

(72) 発明者 今井 亨

東京都世田谷区砦一丁目10番11号 日本放
送協会 放送技術研究所内

(72) 発明者 安藤 彰男

東京都世田谷区砦一丁目10番11号 日本放
送協会 放送技術研究所内

(74) 代理人 100077481

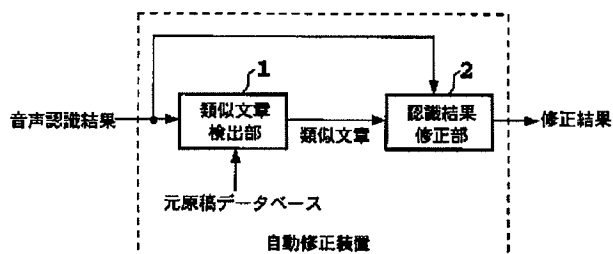
弁理士 谷 義一 (外3名)

(54) 【発明の名称】 音声認識結果の修正方法および装置

(57) 【要約】

【課題】 音声認識結果の修正労力を低減する。

【解決手段】 放送前に作成されたニュース元原稿と、
放送に使用された原稿についての音声認識結果とを比較
し、不一致部分を検出する。ニュース元原稿の不一致部
分を音声認識結果の不一致部分と置換する処理を認識結
果修正部2で行う。



【特許請求の範囲】

【請求項 1】 最終決定した第 1 原稿について発声された音声を生声認識装置により音声認識し、その音声認識結果を修正する音声認識結果の修正方法において、最終決定前の第 2 原稿と前記音声認識結果とを情報処理装置により比較することにより前記第 2 原稿および前記音声認識結果の不一致部分を検出し、当該検出された第 2 原稿側の不一致部分を音声認識結果の不一致部分と置換することにより音声認識結果を修正することを特徴とする音声認識結果の修正方法。

【請求項 2】 請求項 1 に記載の音声認識結果の修正方法において、前記情報処理装置はさらに前記不一致部分の長さを計数し、その計数結果が一定長さよりも小さい場合に前記置換を行うことを特徴とする音声認識結果の修正方法。

【請求項 3】 請求項 1 に記載の音声認識結果の修正方法において、前記第 2 原稿と前記音声認識結果との比較に、音響スコアを使用することを特徴とする音声認識結果の修正方法。

【請求項 4】 請求項 1 に記載の音声認識結果の修正方法において、前記第 2 原稿と前記音声認識結果との比較に、言語スコアを使用することを特徴とする音声認識結果の修正方法。

【請求項 5】 最終決定した第 1 原稿について発声された音声を生声認識装置により音声認識し、その音声認識結果を修正する音声認識結果の修正方法において、最終決定前の複数の第 2 原稿がデータベースの形態で情報処理装置に記憶されており、前記音声認識結果を前記情報処理装置に入力し、前記情報処理装置は、当該入力された音声認識結果と最も類似する第 2 原稿を前記データベースから取り出すことを特徴とする音声認識結果の修正方法。

【請求項 6】 請求項 5 に記載の音声認識結果の修正方法において、最も類似する第 2 原稿の取り出しのために動的計画法を使用することを特徴とする音声認識結果の修正方法。

【請求項 7】 請求項 6 に記載の音声認識結果の修正方法において、前記動的計画法では単語単位で前記音声認識結果と前記データベース上の第 2 原稿との照合を行うことを特徴とする音声認識結果の修正方法。

【請求項 8】 最終決定した原稿について発声された音声を生声認識装置により音声認識し、その音声認識結果を修正する音声認識結果の修正装置において、最終決定前の第 2 原稿と前記音声認識結果とを比較することにより前記第 2 原稿および前記音声認識結果の不一致部分を検出する手段と、当該検出された第 2 原稿側の不一致部分を音声認識結果

の不一致部分と置換することにより音声認識結果を修正する手段とを具えたことを特徴とする音声認識結果の修正装置。

【請求項 9】 請求項 8 に記載の音声認識結果の修正装置において、前記情報処理装置はさらに前記不一致部分の長さを計数する手段を有し、その計数結果が一定長さよりも小さい場合に前記置換を行うことを特徴とする音声認識結果の修正装置。

【請求項 10】 請求項 8 に記載の音声認識結果の修正装置において、前記不一致部分を検出する手段は、前記第 2 原稿と前記音声認識結果との比較に、音響スコアを使用することを特徴とする音声認識結果の修正装置。

【請求項 11】 請求項 8 に記載の音声認識結果の修正装置において、前記不一致部分を検出する手段は、前記第 2 原稿と前記音声認識装置との比較に、言語スコアを使用することを特徴とする音声認識結果の修正装置。

【請求項 12】 最終決定した第 1 原稿について発声された音声を生声認識装置により音声認識し、その音声認識結果を修正する音声認識結果の修正装置において、最終決定前の複数の第 2 原稿をデータベースの形態で記憶する手段と、前記音声認識結果を入力する手段と、当該入力された音声認識結果と最も類似する第 2 原稿を前記データベースから取り出す手段とを具えたことを特徴とする音声認識結果の修正装置。

【請求項 13】 請求項 12 に記載の音声認識結果の修正装置において、最も類似する原稿の取り出しのために動的計画法を使用することを特徴とする音声認識結果の修正装置。

【請求項 14】 請求項 13 に記載の音声認識結果の修正装置において、前記動的計画法では単語単位で前記音声認識結果と前記データベース上の第 2 原稿との照合を行うことを特徴とする音声認識結果の修正装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、音声認識結果を修正するための音声認識結果の自動修正方法および装置に関する。

【0002】

【従来の技術】 従来、音声認識装置により音声認識された結果（文字列）を修正する場合は、音声認識結果を表示装置の画面に表示させたり、印刷装置により印刷して、表示または印刷結果をチェック者が目視により確認する。

【0003】 その後、コンピュータのワープロ（文書編集）機能を使用して、音声認識結果を文書編集（文字列の変更、削除、挿入等）を行い、音声認識結果を修正する。

【0004】

【発明が解決しようとする課題】チェック者が音声認識結果を修正する場合、目視確認とキーボード等を利用した文書編集操作を伴うので、チェック者にとっては非常に煩雑であり、大変な労力を要する。

【0005】そこで、本発明の第1の目的は、上述の点に鑑みて、音声認識結果を自動修正することのできる音声認識結果の修正方法および装置を提供することにある。

【0006】本発明の第2の目的は、音声認識結果のチェックに関連するチェック者の操作労力を軽減することの可能な音声認識結果の修正方法および装置を提供することにある。

【0007】

【課題を解決するための手段】上記第1の目的を達成するために、請求項1の発明は、最終決定した第1原稿について発声された音声声を音声認識装置により音声認識し、その音声認識結果を修正する音声認識結果の修正方法において、最終決定前の第2原稿と前記音声認識結果とを情報処理装置により比較することにより前記第2原稿および前記音声認識結果の不一致部分を検出し、当該検出された第2原稿側の不一致部分を音声認識結果の不一致部分と置換することにより音声認識結果を修正することを特徴とする。

【0008】請求項2の発明は、請求項1に記載の音声認識結果の修正方法において、前記情報処理装置はさらに前記不一致部分の長さを計数し、その計数結果が一定長さよりも小さい場合に前記置換を行うことを特徴とする。

【0009】請求項3の発明は、請求項1に記載の音声認識結果の修正方法において、前記第2原稿と前記音声認識結果との比較に、音響スコアを使用することを特徴とする。

【0010】請求項4の発明は、請求項1に記載の音声認識結果の修正方法において、前記第2原稿と前記音声認識結果との比較に、言語スコアを使用することを特徴とする。

【0011】上記第2の目的を達成するために、請求項5の発明は、最終決定した第1原稿について発声された音声声を音声認識装置により音声認識し、その音声認識結果を修正する音声認識結果の修正方法において、最終決定前の複数の第2原稿がデータベースの形態で情報処理装置に記憶されており、前記音声認識結果を前記情報処理装置に入力し、前記情報処理装置は、当該入力された音声認識結果と最も類似する第2原稿を前記データベースから取り出すことを特徴とする。

【0012】請求項6に記載の発明は、請求項5に記載の音声認識結果の修正方法において、最も類似する第2原稿の取り出しのために動的計画法を使用することを特徴とする。

【0013】請求項7に記載の発明は、請求項6に記載の音声認識結果の修正方法において、前記動的計画法では単語単位で前記音声認識結果と前記データベース上の第2原稿との照合を行うことを特徴とする。

【0014】請求項8に記載の発明は、最終決定した原稿について発声された音声声を音声認識装置により音声認識し、その音声認識結果を修正する音声認識結果の修正装置において、最終決定前の第2原稿と前記音声認識結果とを比較することにより前記第2原稿および前記音声認識結果の不一致部分を検出する手段と、当該検出された第2原稿側の不一致部分を音声認識結果の不一致部分と置換することにより音声認識結果を修正する手段とを具えたことを特徴とする。

【0015】請求項9に記載の発明は、請求項8に記載の音声認識結果の修正装置において、前記情報処理装置はさらに前記不一致部分の長さを計数する手段を有し、その計数結果が一定長さよりも小さい場合に前記置換を行うことを特徴とする。

【0016】請求項10に記載の発明は、請求項8に記載の音声認識結果の修正装置において、前記不一致部分を検出する手段は、前記第2原稿と前記音声認識結果との比較に、音響スコアを使用することを特徴とする。

【0017】請求項11に記載の発明は、請求項8に記載の音声認識結果の修正装置において、前記不一致部分を検出する手段は、前記第2原稿と前記音声認識結果との比較に、言語スコアを使用することを特徴とする。

【0018】請求項12に記載の発明は、最終決定した第1原稿について発声された音声声を音声認識装置により音声認識し、その音声認識結果を修正する音声認識結果の修正装置において、最終決定前の複数の第2原稿をデータベースの形態で記憶する手段と、前記音声認識結果を入力する手段と、当該入力された音声認識結果と最も類似する第2原稿を前記データベースから取り出す手段とを具えたことを特徴とする。

【0019】請求項13に記載の発明は、請求項12に記載の音声認識結果の修正装置において、最も類似する原稿の取り出しのために動的計画法を使用することを特徴とする。

【0020】請求項14に記載の発明は、請求項13に記載の音声認識結果の修正装置において、前記動的計画法では単語単位で前記音声認識結果と前記データベース上の第2原稿との照合を行うことを特徴とする。

【0021】

【発明の実施の形態】以下、図面を参照して本発明の実施形態を詳細に説明する。

【0022】最初に本発明を適用した音声認識結果の自動修正方法について説明する。

【0023】(a)放送における原稿
ニュースでは、アナウンサーが放送中に直接読む、アナウンサー原稿と、あらかじめ放送前に、記者が入力して

電子化された記者原稿がある。ニュース項目の1つに対して、いろいろな切り口で書かれた、複数の記者原稿が存在し、それをアナウンサーが読むために紙に印刷し、放送に入る直前まで、アナウンサー本人やディレクターらによって、手書きで加筆や訂正が行われている。

【0024】自動字幕化のために、直接記者原稿を用いることができないのは、放送に入る直前まで行われる手書きの修正に対して、記者原稿を訂正する時間的な余裕がなく、また、記者原稿にないアドリブにも対応できないからである。

【0025】ニュースでのアナウンサーの発声1文と記者原稿との対応関係は、以下のように大別できる。

【0026】(1) アナウンサーの発声1文に対して、記者原稿の1文が対応する場合

(2) アナウンサーの発声1文に対して、記者原稿の複数の文が対応する場合

(3) アナウンサーの発声1文に対して、記者原稿が存在しない場合

上記対応の割合は約7:1:2になっており、まったく同じではないにしろ、8割は元の記者原稿が存在してい

$$T(W_i, W_j) = \begin{cases} S - \alpha \cdot (|j - i - 1|); & S - \alpha \cdot (|j - i - 1|) \geq \alpha, \\ \alpha; & S - \alpha \cdot (|j - i - 1|) < \alpha \end{cases} \quad (i \leq j)$$

【0031】とした。但し、 S は遷移スコアの最大値、 α は距離によって減少する値である。

【0032】単語単位の類似度は、文字単位の類似度から求める。単語AとBの類似度 $W(A, B)$ は、

【0033】

【数2】

$$W(A, B) = \max_{\text{全てのパス上で}} \{m \cdot H + n \cdot K\}$$

【0034】とした。但し、 m , n はそれぞれ単語AB間で、 m は平仮名のマッチングした文字数、 n は平仮名以外の文字のマッチングした文字数である。定数 H , K はそれぞれ、平仮名のマッチングスコア、平仮名以外の文字のマッチングスコアである。なお、文字間の遷移についてはスコアを与えていない。

【0035】対応する文が存在するかしないかの判定には、しきい値を用いた。認識結果の単語数を N とすると、しきい値 θ は

【0036】

【数3】 $\theta = N \cdot (H + K) / 2$

とした。

【0037】(b)-2 対応する原稿による認識結果の修正

認識結果の修正法は、マッチングした単語の間を原稿の単語で順次埋めていく形をとった。これにより、未知語として認識できなかった単語も認識結果に加えることが可能になる。ただし、原稿内に実際には読まれていない

る。

【0027】(b) 認識結果の修正法

ニュース音声認識結果の修正は、以下の二段階で行う。まず第一段階として、ニュース音声の認識結果を利用して、この元になったと思われる記者原稿の文を検出する。第二段階として、検出された記者原稿の文を利用して、認識結果の単語の入れ替えを行う。

【0028】(b)-1 認識結果に対応する元原稿の検出

認識結果に対応する記者原稿を検出するために、文単位のDPマッチングを行う。同じ放送日の記者原稿をすべて形態素解析して単語に分割し、最も認識結果との類似度の高い文を、対応する元の文とみなすことにする。

【0029】文単位の類似度を、単語単位の類似度と、単語間の遷移スコアの総和が最大のものとする。文 $W_1 W_2 W_3 \dots W_n$ における、単語 W_i から W_j への遷移スコア T は、

【0030】

【数1】

単語の挿入を防ぐために、置換する両者の単語数の差が

【0038】

【数4】 $|C_1 - C_2| < L$

である時のみ置換を行った。ここで、置換する認識結果の単語数を C_1 、元原稿の単語数を C_2 とした。また、認識結果と記者原稿のDPマッチングでは始点と終点をフリーとしたので、認識結果の文頭と文末の単語がマッチングしなかった時には、認識結果をそのまま使用した。

【0039】(c) 実験

(c)-1 実験条件

実験には、平成8年6月4日のある放送局のニュースの音声認識結果53文(総単語数1,804)(小林ほか、“ニュース音声認識システムの検討”音講論集, 3-1-9, pp. 103-104(1997.9)と、同日の記者原稿データベースを使用した。修正前の認識結果は、単語正解精度67.2%である。

【0040】(c)-2 実験結果

まず、文字のマッチングスコアを、設定するために H と K の重みを変化させて対応する文の検出に最適な値を調べた。ただし、単語間の遷移スコア S については、平仮名のマッチングスコア H と同値となるようにし、 L は5単語以内の差は許すとして、 $L=5$ とした。その結果を表1に示す。

【0041】

【表1】

HとKの対応の調査

HとKの重みの比 (K/H)	1 倍	2 倍	3 倍	4 倍	5 倍
原稿の検出正解率 (%)	73.6	73.6	90.6	83.0	83.0

【0042】平仮名以外（漢字、カタカナなど）の文字と平仮名ではマッチングスコアには3倍（ $K = 3 * H$ ）の差をつけた場合が今回はよかった。表2に詳しく示す。

【0043】その結果、認識結果と原稿との対応関係は元原稿が無い場合も含めて、90.6%正しく検出できた。文字のマッチングスコアを変化させてみると、今回のように漢字やカタカナにある程度、重点を置いた場合

が有効であることがわかった。検出に失敗したものは、ほとんど認識時の正解精度の低いものであった。また、認識結果が短いほど、誤った文を検出しやすい。検出誤りの内、1文は原稿中の複数の文に対応していたため、対応文が無いと判定された。

【0044】

【表2】

対応文の検出結果

元原稿が存在するもの		40文	全体の75.5%
	対応する文を検出	36文	原稿ありの90.0%
	対応しない文を検出	0文	原稿ありの0.0%
	文を検出しなかった	4文	原稿ありの10.0%
元原稿が存在しないもの		13文	全体の24.5%
	対応しない文を検出	1文	原稿なしの7.7%
	文を検出しなかった	12文	原稿なしの92.3%
計		53文	
		検出の正解率 90.6%	

【0045】次に、提案手法で音声認識結果を修正した。その結果、表3に示すように、14.3%の改善がみられた。

【0046】中には、認識結果に含まれていなかった未知語が加えられ、改善された場合が見られた。もともとの認識結果とほとんど認識精度が変わらないものでも、原稿の単語に入れ替えられているために、より日本語の文として意味の通るものになっている。しかし、記者原稿の表記の違いや誤字については無条件で修正されてしまったり、原稿の前後の順番を入れ替えて作られたアナウンサー原稿を修正できない、などの問題点がある。

【0047】

【表3】

修正結果

	単語正解精度
修正前	67.2%
修正後	81.5%

【0048】(d) まとめ

記者原稿を用いて認識結果を修正することで、認識精度の改善が見られた。今回は、文ごとに検出したので複数文にわたる認識結果には対応していない。また、単に原稿の文字について比較し修正を行ったが、精度向上のために、音響的なスコアを利用したり、認識結果の単語間のbigramなどを使用して修正を行なうことができる。

【0049】次に本発明を適用した音声認識結果の自動

修正装置について説明する。

【0050】（実施形態1）図1は本発明実施形態1のシステム構成を示す。なお、自動修正装置はコンピュータ等の情報処理装置が後述の処理を規定したプログラムを実行することによっても実現することができる。図1において、自動修正装置は、元原稿データベース中で音声認識結果に最も類似した文章を検出する「類似文章検出部」1と、類似文章によって音声認識結果を修正する「認識結果修正部」2を有する。

【0051】次に、類似文章検出部1について、図2を参照して説明する。類似文章検出部1は、まず照合スコア計算部11において、あらかじめ形態素解析によって単語単位に分割された自動修正装置内の元原稿データベ

$$T(W_i, W_j) = \begin{cases} S - \alpha \cdot (|j - i - 1|); & S - \alpha \cdot (|j - i - 1|) \geq \alpha, \\ \alpha; & S - \alpha \cdot (|j - i - 1|) < \alpha \text{ となる。} \end{cases} \quad (i \leq j)$$

【0053】但し、 S は遷移スコアの最大値、 α は距離によって減少する値である。また、単語単位の照合スコアは、単語AとBの照合スコア $W(A, B)$ は、

【0054】

【数6】

$$W(A, B) = \max_{\text{全てのパス上で}} \{m \cdot H + n \cdot K\}$$

【0055】となる。但し m 、 n はそれぞれ単語AB間で、 m は平仮名のマッチングした文字数、 n は平仮名以外の文字の照合が一致した文字数である。定数 H 、 K はそれぞれ、平仮名の照合スコア、平仮名以外の文字の照合スコアである。

【0056】類似文章判定部12では、最大の照合スコアを示す元原稿について、音声認識結果の単語数に比例した閾値以上であれば、その元原稿を音声認識結果の類似文章とみなし、そうでなければ類似文章が存在しないものとみなす。たとえば、実施形態の場合、認識結果の単語数を N とすると、しきい値 θ は、 $\theta = N \cdot (H + K) / 2$ と表現できる。

【0057】次に、認識結果修正部2について、図3を参照して説明する。類似文章検出部1で類似文章が存在すると判定された場合、単語不一致区間検出部21において、音声認識結果と類似文章の単語が一致しない区間を検出する。次に置換判定部22において、不一致区間の音声認識結果と類似文章の各々の単語数の差がある一定数以下である場合、音声認識結果の単語を類似文章の単語で置換する。たとえば、置換する両者の単語数の差は置換する認識結果の単語数を $C1$ 、元原稿の単語数を $C2$ とする場合、 $|C1 - C2| < L$ と表現することができる。

【0058】以上の処理をコンピュータで実行するためのプログラムの処理内容を図4に参考のために示しておく。説明が上述の説明と重複する箇所についてはその説

明を用いて、音声認識結果と元原稿の各文章を単語単位に動的計画法（DPとも呼ばれる）で照合し、照合スコアを求める。この際、照合スコアは平仮名と平仮名以外の文字について、照合が一致した場合のスコアに数倍の差をつけ、また、単語間の遷移スコアは照合距離に反比例するものとする。たとえばこの実施形態において、動的計画法の照合スコアを P 、各単語どうしのマッチングの照合スコアを W 、単語間の遷移スコアを T とすると、 $P = \sum (W + T)$ と表現できる。また、文 $W1W2W3 \dots Wn$ における、単語 Wi から Wj への遷移スコア T は、

【0052】

【数5】

明を簡単に留める。

【0059】修正対象の音声認識結果は、通信、フロッピーディスク、キーボード等の入力装置から入力され、コンピュータ内のメモリに格納される（ステップS10）。比較の対象となる原稿はこの実施形態ではコンピュータ内にデータベースから与えられるが、予め対応する原稿が判明している場合は通信、フロッピーディスク、キーボード等の入力装置から与えてもよい。本実施形態では音声認識結果と対応する原稿をも自動検出する。

【0060】最初にデータベースから取り出された原稿と修正対象の音声の認識結果との照合を上述の動的計画法で行う（ステップS20～S40）。動的計画法を使用すると、単語同士の類似度を計算するだけでなく、連続する単語列についての類似度についても計算（各単語の類似度を重み付けして加算、数1、数2式参照）するので、一致する単語が多く、かつ、その単語の並びの順序が一致する場合ほど照合スコアは高くなる。

【0061】得られた照合スコアを認識結果の単語数から定めた閾値と比較し、データベースから取り出した原稿が音声認識結果と対応するか否かを判定する（ステップS50）。

【0062】閾値よりも照合スコアが低い場合には、データベースから取り出した原稿は、音声認識結果とは対応しないので（ステップS50でNO判定）、次の原稿と音声認識結果を照合すべく、ステップS90を経由して、ステップS20へ手順を戻す。

【0063】これにより、次の原稿がコンピュータにより（正確にはCPU）データベースから取り出されて、照合が行なわれる。このような処理を繰り返し、対応する原稿がデータベースから取り出されると、その原稿の照合スコアは閾値よりも高くなる（ステップS50がYES判定）、これにより手順は、ステップS60に進

み、音声認識結果と、現在の原稿の文章とを比較することにより不一致の単語列部分を検出する。

【0064】また、不一致の単語部分の単語数を計数する。この単語数が一定数以下の場合には、不一致部分の原稿側の単語列で、音声認識側の不一致部分の文字列を置換する（ステップS70→S80）。

【0065】一方上記不一致部分の単語数が一定値よりも多い場合には、音声認識結果に誤認識があるのではなく、元原稿が修正されて、放送されたと判定して、置換処理は行わず、別の原稿をデータベースから取り出す（ステップS70→S90→S20）。

【0066】なお、上述のステップS20～S90の処理を繰り返し実行しても照合スコアが閾値以上で、かつ、不一致の単語数が一定数以下となるデータベース上の原稿が見つからなかった場合には、手順を終了する（ステップS90のYES判定→終了）。

【0067】以上の処理を実行することにより、音声認識結果の自動修正が可能となるとともに、単語あるいは文字の修正に用いる原稿をデータベースから自動的に検出することも可能となる。

【0068】（実施形態2）図5には、認識結果修正部2のもう一つの実施形態を示す。これは、図3の認識結果修正部に「音響スコアおよび言語スコア計算部」32を加え、テキストの表記を単純に置き換えるのではなく、修正後のテキストが音響的にも言語的にもより正しいかどうかを、認識対象音声を用いて詳細に判定するのである。まず、単語不一致区間検出部31において、音声認識結果と類似文章の単語不一致区間を求める。次に音響スコアおよび言語スコア計算部32において、音声認識結果および類似文章の単語不一致区間に存在するすべての単語を用いて単語ネットワークを作成し、認識対象音声の音響スコアおよび／または言語スコアを計算する。最後に置換判定部33において、音響スコアと言語スコアの重み付け和が最大となる単語列で、音声認識結果の単語不一致区間を置換する。

【0069】なお音響スコアまたは言語スコアのいずれか一方のみを使用して単語不一致区間を置換することもできる。

【0070】このための実際例としては、実施形態1のステップS70の単語数によって置換を判定している部分（置換判定部22）を、再度、認識結果と類似文章の不一致区間の単語のみで、対応する音声区間を音声認識装置にかけ音響スコアと言語スコアを計算したものと、もともとの認識結果が認識時に得た不一致区間の音響スコアと言語スコアを比較して置換するもの（音響スコアおよび言語スコア計算部32・置換判定部33）に置き換えたものである。ここで、用いている音響スコアとは連続音声認識装置で使用しているHMMと不一致区間に対応する音声との計算したスコアであり、また、言語スコアとは各7年間の原稿を利用し音声認識装置用にあ

はじめ学習した、単語連鎖確率を統計的に計算したbigram・trigramのことを示す。

【0071】認識結果の類似原稿との不一致区間の、もともとの認識時にえた音響スコアと言語スコアの合計をPとし、認識結果と類似文章の不一致区間の単語のみを使用して不一致区間に対応する音声に対して音声認識装置が出力した音響スコアと言語スコアの合計をPr（音響スコアおよび言語スコア計算部）とすると、 $P < Pr$ の時のみ（置換判定部33）不一致区間に対応する音声に対して行った音声認識装置の出力結果を不一致区間の正解とみなし置換する。

【0072】本実施形態の他に次の形態を実施できる。

【0073】1）上述の実施形態では音声認識を自動修正する例を示したが、チェック者が音声認識結果と類似する原稿をデータベースから取り出したいときには上述の動的計画法を使用した類似度の計算を行って、最も類似する原稿を検出し、その原稿をプリンタにより印刷したり、表示器に表示させることもできる。この場合には、チェック者は、表示、あるいは印刷出力を見て音声認識結果を手動で修正するが、チェックの基となる原稿を自動的にデータベースから取り出すことができるので、原稿の検索にかかわる操作から開放される。

【0074】2）さらに上述した実施形態において、修正の基となった原稿および音声認識結果をともに表示させ、音声認識結果の修正された個所を他の部分と異なる色で表示すると、チェック者は修正個所を目視確認することができる。

【0075】3）上述の実施形態で述べたコンピュータにはパソコン、ワークステーション、ワープロ等各種の情報処理装置を使用することができる。

【0076】

【発明の効果】以上、説明したように、請求項1、8の発明によれば、最終決定前の第2原稿は最終決定した原稿と共通部分が多い点、および最終決定した第1原稿を放送するまでの間に第2原稿を電子化（文書として情報処理装置に入力すること）している点に着目し、第1原稿の音声認識結果と第2原稿とを比較することで、音声認識結果を速やかに自動修正することができる。

【0077】請求項2、9の発明では不一致部の長さ、たとえば、文字数や単語数などの長さを計数することで、不一致部分が音声の誤認識により発生したもののか、第2原稿から第1原稿を作成する際に加えられた変更かを区別することができるので、誤認識部分を自動検出することができる。

【0078】請求項3、4、10、11の発明では、さらに音響スコアや言語スコアを加味することにより誤認識部分を精度よく検出することができる。

【0079】請求項5～7、12～14の発明は、動的計画法により第2原稿の文章を音声認識結果と比較することで、原稿の類似度を計算することができる点に着目

して、最も音声認識結果と類似する原稿をデータベースから自動的に取り出すことにより、操作者はデータベースの検索に関わる操作から解放される。

【図面の簡単な説明】

【図１】本発明実施形態１の構成を示すブロック図である。

【図２】類似文章検出部１の構成を示すブロック図である。

【図３】認識結果修正部２の構成を示すブロック図である。

【図４】音声認識結果の自動修正処理手順を示すフロー

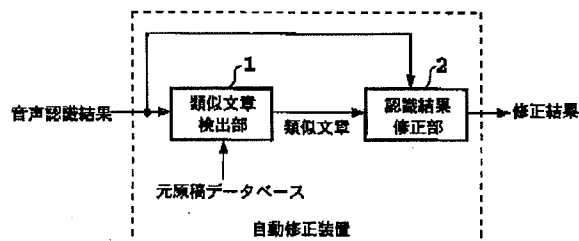
チャートである。

【図５】実施形態２の構成を示すブロック図である。

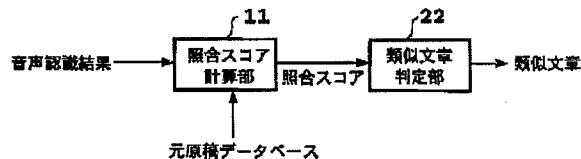
【符号の説明】

- １ 類似文章検出部
- ２ 認識結果修正部
- １１ 照合スコア計算部
- １２ 類似文章判定部
- ２１、３１ 単語不一致区間検出部
- ２２、３３ 置換判定部
- ３２ 置換判定部

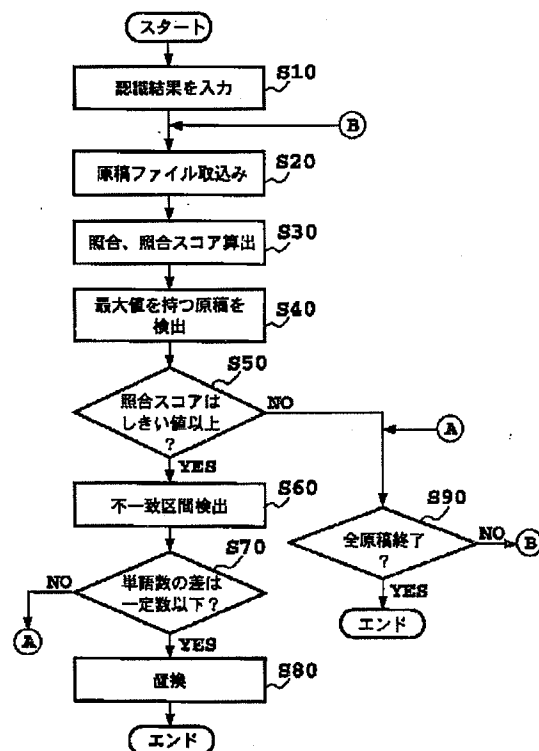
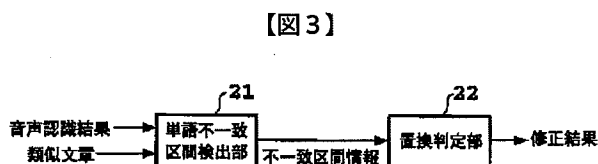
【図１】



【図２】



【図４】



【図 5】

